



MASTER FILE

April 12, 2000

DSSD CENSUS 2000 PROCEDURES AND OPERATIONS MEMORANDUM SERIES #O-7

MEMORANDUM FOR Michael J. Longini
Chief, Decennial Systems and Contracts Management Office

From: Howard Hogan *Howard Hogan*
Chief, Decennial Statistical Studies Division

Prepared By: Kimball Jonas *Kimball Jonas*
Decennial Statistical Studies Division

Subject: Observation of Operational Test and Dry Run of
Group Quarters Forms at the Data Capture Center
in Jeffersonville, IN

I. INTRODUCTION

During the week of February 1 - 4, 2000, I observed a portion of the Operational Test and Dry Run (OTDR) at the Jeffersonville, IN, Data Capture Center (DCC). My observations were focused on the check-in and data capture procedures for Group Quarters (GQ) forms, particularly the short and long Individual Census Reports (ICRs) and the Military Census Reports (MCRs), known as the D-20A, D-20B, and D-21, respectively. I also observed DCC check-in of the short and long Individual Census Questionnaires (ICQs), that is, the D-15A and D-15B, which was also included in the OTDR.

II. OVERVIEW OF DATA CAPTURE PROCEDURES FOR GQ FORMS

A. Layout of the Jeffersonville DCC

The Jeffersonville DCC takes up bays A through E of Building 60 of the Census Bureau's National Processing Center (NPC) in Jeffersonville, IN. In addition, Key-From-Image (KFI) keying for forms captured by the DCC takes place in Building 63G¹ of that facility.

In Building 60, the A bay is dedicated to sorter check-in; forms that arrive in barcoded envelopes that can be machine sorted by form type begin here. Manual check-in (GQ forms are checked in manually) takes place in the B bay; C bay is given over to document preparation and staging for scanning; D bay is dedicated to scanning, and E bay is divided between keying from paper (KFP) and checkout.

¹A specific bay in a particular building is referred to by the conjoining of building number and bay letter, as in this instance: Building 63G is shorthand for bay G of Building 63.

B. Data Capture of GQ Forms

On arrival at the loading dock of B bay, GQ forms will be in envelopes by GQ, with the GQ's Form D-352 (GQ Enumeration Record) cover sheet taped to the outside of the envelope. Envelopes for a number of GQs will have been put in a box at the LCO; the box will contain only GQ forms (ICQs, ICRs, and MCRs).

After arrival at B bay, the boxes of GQ forms are wanded in, using a scanning gun. Then the individual GQs represented in each box are checked in: on the check-in clerk's computer, a screen is called up with blanks for the LCO code, the GQID, the expected number of people, and the actual count. The 'expected' number is the LCO's actual count; the clerk hand-counts the forms to obtain the 'actual' count. The clerk types in the numbers; the computer gives an error message if the GQID fails the MAD 97 check digit test. The clerk checks the GQIDs on the individual questionnaires against the GQID on the D-352, and corrects any that are in error.

When clerks have completed check-in of the GQ forms, what happens next depends on form type. The ICQs and Puerto Rico ICRs, which are keyed outside the DCC, are put back in their original envelope, and the envelopes placed in a bin to be transported to the check-in area for keying in Building 64-A. The MCRs and all other ICRs, which are scanned in the DCC, are sorted by form type, placing them in different trays on a metal cart. A pre-printed, barcoded batch header sheet is placed on each tray. An 8 ½ x 11 tag is placed in a transparent plastic sleeve on the cart, indicating the cart's destination. Control clerks are responsible for moving the carts from one bay to the next.

The batch sheets have a bar code on one side, and a preprinted Batch Activity Record on the other, with blanks for the employee number, initials, and date of each of several operations: doc prep, scanning, KFP-Enter, KFP-Verify, checkout clerk, checkout analyst, rebatch, To KFP-Enter, To KFP-Verify, and To Scanning.

From manual check-in, the GQ forms remaining in the DCC go to Document Preparation (doc prep). For MCRs and long form ICRs, doc prep consists of 'relaxing the creases' in the forms, i.e. physically manipulating the forms so that the folds lie flat, and laying the forms flat, all facing the same way, in a cardboard tray long enough to accommodate them. For the short form ICRs, doc prep consists of stacking the forms neatly in a cardboard tray, all facing the same way.

After doc prep, the trays are put on another cart, and sent to scanning. During my observation, the scanner was able to scan the ICR short forms quite rapidly; it was a bit slower with the ICR long forms and the MCRs, but was able to successfully capture the images without problems. When a long form came out of the scanner, it frequently tended to 'catch' on the folds between the pages of the previous form. But since the image had already been successfully captured, that was not a particularly serious problem.

After successful scanning of the image, the image would go to 'data lift,' as the

OCR/OMR data recognition process is known. This step takes place out of sight, inside the server. If the computer has insufficient confidence that it has interpreted the image of a field correctly, it will send the image to KFI in Building 63G, where the images to be keyed are queued to the keyers on duty. The image of the field on the original questionnaire, including a generous area surrounding the field, appears on the keyer's screen, along with a computerized field for the keyer to type in her interpretation of the respondent's handwritten answer.

Once data lift and KFI are completed, the batch goes to checkout. A checkout clerk records the batch in the check-out log book. Each form type has its own page(s) in the log book; the clerk records her name and the batch ID. To commence check-out, she scans the bar code on the batch header sheet, then pulls the processing ID on each form under a scanning gun resting on a stand. For each ID scanned, the clerk's screen lights up with a green or red light, indicating a successful data capture (green) or capture problems (red). The batch then goes to a check-out analyst.

The check-out analyst wands the bar code of the batch header sheet. If every form in the batch has been successfully captured, then a "tray is ready for storage" message appears on the analyst's screen, and she directs it there. If some forms in the batch have not successfully been captured (including forms that inadvertently were not scanned at all), then the analyst is responsible for determining the correct disposition of those forms (i.e., KFP, imaging, manual check-in, etc.). The screen will give a recommended disposition, such as "Send to Imaging or KFP"; the analyst can make the decision based on the condition of the form and the reason for the failure that appears on the analyst's screen. The analyst separates those forms out and puts them in new batches to be re-scanned, keyed from paper, etc., as appropriate. (If forms were scanned but are missing at check-out, no error message appears.) The system then allows the analyst to send the remainder of the batch, consisting of successfully captured forms, to storage.

III. SPECIFIC OBSERVATIONS

With the exception of the arrival of forms on the loading dock, and the KFI keying, all of the foregoing is from actual observation. Some further particulars:

- GQ check-in went quite smoothly, once the clerks became used to the distinct procedures for GQ forms. I was impressed by their thoroughness in dealing with mundane tasks such as counting and re-counting the forms from a GQ.
- On the first batch of D-20B forms scanned, six consecutive forms received a "critical area missing" message at the scanner analyst's station, a message usually indicative of a torn or folded form. None of the forms' images had any visible problems, and the reason for the error message was never found. The scanned images were successfully captured, and the error did not repeat itself.
- The time it took for a batch to arrive at KFI after being scanned was fundamentally unpredictable. As a result, I was unable to observe KFI of GQ forms.
- On the D-20A form, a second barcode is positioned very close to the processing

ID bar code. It was easy to accidentally scan that barcode while scanning the processing ID; checkout clerks had to be especially careful to avoid that error. There was no similar problem with the D-20B forms.

IV. CONCLUSIONS

Nothing in the dry run suggested the possibility of any serious problems with capturing the GQ forms. I came away with a great deal of respect for the intelligence and competence of the personnel in charge of data capture at the Jeffersonville DCC, and I expect them to be capable of handling any problems that might arise with data capture of GQ forms during Census 2000.

cc:

DSSD Census 2000 Procedures and Operations Memorandum Series List

L. Affleck	(DSCMO)	J. Lothar	(GEO-ALRB)
D. Burkhead		E. Jaramillo	(CLO)
J. Dickens		S. Rolark	
G. McLaughlin		E. Byerly	(POP)
D. Stoudt		A. Smith	
J. Stovall		D. Smith	
J. Benetti	(FLD)	L. Kehm	
S. Schoch		L. Moyer	(DMD)
M. Weiler		D. Dalzell	(HHES)
J. Dawson	(TMO)	T. McNally	(DSSD)
K. Wyatt			
C. Falkenstein	(GEO-MAF))		
J. McKay			